

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Wolf, Marcia
Cassels, Frederick
Boedeker, Edgar
- (ii) TITLE OF INVENTION: Transformed Bacteria Producing CS6
Antigens and Vaccines
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Glenna Hendricks
 - (B) STREET: P.O. Box 2509
 - (C) CITY: Fairfax
 - (D) STATE: VA
 - (E) COUNTRY: USA
 - (F) ZIP: 22031
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 09/479,877
 - (B) FILING DATE: 10-JAN-2000
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Hendricks, Glenna
 - (B) REGISTRATION NUMBER: 32,535
 - (C) REFERENCE/DOCKET NUMBER: army 09/479,877
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (703) 425-8405
 - (B) TELEFAX: (703) 425-8406

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4876 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: E. coli

(B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAGCTTGTA CCAGTTGATA AAAATATATC ACGCTGGGAA TGACGTGATG TATATACGGA
60

GCAGCTATGT CGGAACAGAT ATTTTCCTAT CGGTATGCGT TGTGAGTAAG CGTAAAGCCA
120

ATGCTGTCTG TAACTCCTGA TCCTTGCAGA CTAAATTAGA GCTCCTTCTA AATTAGACGG
180

ATGGATAAAC CTACAGACTG GCGCTCTGGG TCTCGCCGGA TATTTTCTAA TGAATTTAAG
240

CTTCATATGG TTGAACTGGC TTCGAAACCA AATGCCAATG TCGCACAACT GGCTCGGGAA
300

CATGGCGTTG ATAACAACCT GATTTTTTAAA TAGCTACGCC TCTGGCAAAG AGAAGGACGT
360

ATTTCTCGTA GAATGCCTCC AACTATTGTA GGCCCTACAG TACCACTGAG GTAGCCTGAA
420

TTTAAAGCCG AAGCGGTCAG AACTGTTCTT GGTGTGAACG TAGCACTCAC CAATAAAAGC
480

ATCAATACGG TGCTCTGTTG ACACATTACG AATGTTATGT ATACAATAAA AATGATTATA
540

GCAATATTAA TGGTGTATA TGAAGAAAAC AATTGGTTTA ATTCTAATTC TTGCTTCATT
600

CGGCAGCCAT GCCAGAACAG AAATAGCGAC TAAAACTTC CCAGTATCAA CGACTATTTT
660

AAAAAGTTTT TTTGCACCTG AACCACGAAT ACAGCCTTCT TTTGGTGAAA ATGTTGAAA
720

GGAAGGAGCT TTATTATTTA GTGTGAACTT AACTGTTCTT GAAAATGTAT CCCAGGTAAC
780

GGTCTACCCT GTTTATGATG AAGATTATGG GTTAGGACGA CTAGTAAATA CCGCTGATGC
840

TTCCCAATCA ATAATCTACC AGATTGTTGA TGAGAAAGGG AAAAAAATGT TAAAAGATCA
900

TGGTGCAGAG GTTACACCTA ATCAACAAAT AACTTTTAAA GCGCTGAATT ATACTAGCGG
960

GGAAAAAAA ATATCTCCTG GAATATATAA CGATCAGGTT ATGGTTGGTT ACTATGTAAA
1020

CTAAATACTG GAAGTATGAT TATGTTGAAA AAAATTATTT CGGCTATTGC ATTAATTGCA
1080

GGAAGTTCCG GAGTGGTAAA TGCAGGAAAC TGGCAATATA AATCTCTGGA TGTAATGTGA
1140

AATATTGAGC AAAATTTTAT TCCAGATATT GATTCCGCTG TTCGTATAAT ACCTGTTAAT
1200

TACGATTCGG ACCCGAAACT GGATTCACAG TTATATACGG TTGAGATGAC GATCCCTGCA
1260

GGTGTAAGCG CAGTTAAAAT CGCACCAACA GATAGTCTGA CATCTTCTGG ACAGCAGATC
1320

GGAAAGCTGG TTAATGTAAA CAATCCAGAT CAAAATATGA ATTATTATAT CAGAAAGGAT
1380

TCTGGCGCTG GTAACTTTAT GGCAGGACAA AAAGGATCCT TTCCTGTCAA AGAGAATACG
1440

TCATACACAT TCTCAGCAAT TTATACTGGT GGCGAATACC CTAATAGCGG ATATTCGTCT
1500

GGTACTTATG CAGGAAATTT GACTGTATCA TTTTACAGCA ATTAAAAAAA GGCCGCATTA
1560

TTGCGGCCAT TGACGATACT GCTAGGCAAA AATATGAAAT CAAAGTTAAT TATACTATTG
1620

ACGTTAGTGC CATTTTCATC TTTTTCACAA GGAAATAATT TTGAAATAAA TAAGACACGA
1680

GTAATTTACT CTGACAGCAC ACCATCAGTT CAAATATCAA ATAATAAAGC ATATCCTTTA
1740

ATTATTCAAA GCAATGTATG GGATGAAAGC AATAATAAAA ATCATGACTT TATAGCAACA
1800

CCACCGATTT TTAAATGGA AAGTGAAAGT CGGAATATAA TAAAAATAAT TAAAACAACT
1860

ATTAATTTGC CGGACTCTCA GGAAAGTATG AGATGGTTAT GTATTGAATC AATGCCACCA
1920

ATAGAAAAAA GTACTAAAAT AAACAGAAAA GAAGGAAGGA CAGACAGTAT TAATATCAGC
1980

ATTCGGGGGT GCATTAACT GATATATCGA CCTGCCAGTG TTCCGTCTCC TGTTTTTAAT
2040

AAATAGTAG AAAAATTAAA ATGGCATAAA AATGGAAAGT ATCTTGTATT AAAAAATAAT
2100

ACACCCTATT ACATTAGCTT TTCTGAGGTT TTTTTTGATT CAGATAAAGT AAACAATGCA
2160

AAAGATATTT TATATGTAAA ACCATACTCA GAGAAGAAAA TAGATATCAG CAACAGAATA
2220

ATAAAAAAAA TCAAATGGGC TATGATTGAT GATGCTGGCG CAAAAACAAA ACTTTATGAA
2280

TCAATTTTAT AAAAAATCTC ATTACAGTAT AAAAAACAT CAGATTACAG GCTTGCTTTT
2340

TTTGCTATTT ATATATCCTT TCTCAACCTC ATATGGAAAT GAACAATTTA GTTTTGACTC
2400

ACGATTCCTA CCATCAGGTT ATAATTACTC TTAAATAGT AACTTACCTC CTGAAGGTGA
2460

GTATCTGGTT GATATTTATA TTAACAAAAT AAAAAAGGAG TCCGCGATTA TTCCTTTTTA
2520

TATAAAAGGA AATAAACTTG TACCATGTTT ATCAAAAGAA AAAATTTTCAT CTTTGGGTAT
2580

CAACATTAAT AATAACGACA ACACAGAGTG TGTAAGGAGC AGTAAGGCAG GTATTAGTAA
2640

TATCAGCTTT GAGTTTAGCT CTCTTCGTTT GTTTATTGCT GTACCGAAAA ATCTTCTGTC
2700

TGAGATTGAT AAAATATCAT CAAAGGATAT AGATAACGGG ATTCATGCTT TATTTTTTAA
2760

TTATCAAGTA AATACAAGGC TAGCCAATAA TAAAAATCGT TATGATTACA TTTCTGTTTC
2820

ACCAAATATA AATTATTTTT CATGGCGGTT GCGTAATCTT TTTGAATTTA ACCAAAACAA
2880

CGATGAAAAA ACATGGGAAA GAACTACAC TTATCTAGAA AAAAGTTTTT ATGATAAAAA
2940

GCTAAACTTA GTCGTTGGTG AAAGTTATAC GAATTCAAAT GTTTATAATA ACTACTCTTT
3000

TACTGGTATT TCAGTTTCTA CAGATACAGA TATGTATACG CCAAGTGAAA TCGATTATAC
3060

ACCAGAAATT CATGGAGTGG CTGATTCAGA CTCTCAGATT ATTGTCAGGC AAGGCAACAC
3120

CATTATCATT AATGAAAGTG TTCCAGCCGG ACCGTTCTCA TTTCCAATAA CCAATCTCAT
3180

GTATACTGGG GGGCAACTTA ATGTGGAGAT AACAGATATT TATGGAAATA AAAACAATA
3240

TACTGTCAAT AATTCCTCTC TTCCTGTTAT GAGAAAAGCG GGACTAATGG TATATAATTT
3300

TATATCTGGG AAATTAACAA AAAAAAATAG TGAGGATGGT GATTTTTTTA CTCAAGGTGA
3360

TATTAACTAC GGTACTCACT ATAACAGCAC ACTATTCGGT GGATATCAGT TTAGTAAAAA
3420

TTATTTTAAAC TTATCTACTG GTATAGGCAC TGATCTGGGA TTTTCTGGAG CATGGCTACT
3480

ACACGTTAGC AGAAGTAATT TTAAGAATAA AAATGGATAT AATATTAATC TACAACAAAA
3540

CACTCAGTTA AGACCATTCA ATGCCGGGGT TAATTTTCGAT TACGCATACA GAAAAAAAAG
3600

GTATGTGGAA CTTTCCGACA TTGGCTGGCA TGGTAATTTA TATAATCAAC TTAAAAATAG
3660

TTTTTCTTTA TCCTTGTCOA AATCATTGAA TAAATACGGA AATTTCTCAC TTGATTATAA
3720

CAAAATGAAA TACTGGGATA ATGCGTATGA TAGTAACTCA ATGTCGATTC GTTATTTTTT
3780

TAAATTCATG CGAGCAATGA TTACAACAAA TTGTTCTTTA AATAAATATC AATCTTATGA
3840

AAAAAAAGAT AAAAGATTTA GTATTAATAT ATCATTGCCT TTAACCAAAG ATTACGGGCA
3900

CATATCTTCA AACTATTCAT TTTCCAATGC AAATACAGGA ACGGCAACCA GTTCTGTAGG
3960

CTTAAACGGT AGTTTTTTTA ATGACGCAAG ATTAAACTGG AACATTCAGC AGAACAGAAC
4020

GACCCGTAAC AATGGATATA CTGATAATAC CAGTTACATA GCAACCAGCT ATGCCTCTCC
4080

CTATGGCGTT TTTACTGGTT CATATTCAGG ATCGAACAAG TATTCAAGCC AGTTTTATTC
4140

TGCATCGGGA GGTATTGTTT TGCATAGCGA TGGCGTAGCT TTTACTCAA AAGCCGGAGA
4200

TACCTCTGCT CTTGTCCGTA TTGATAATAT TTCTGATATA AAAATTGGTA AACTCCTGG
4260

TGTTTATACT GGGTATAATG GTTTTGCTTT AATTCCTCAT CTCAGCCGT TCAAAAAAA
4320

CACCATTTTA ATTAATGATA AAGGAATTCC AGACGGTATT ACTCTTGCTA ATATAAAAA
4380

ACAAGTTATC CCATCACGAG GAGCTATTGT TAAAGTAAA TTTGATGCTA AAAAAGGCAA
4440

TGACATTTTG TTTAAGCTTA CAACTAAAGA TGGAAAAACG CCCCATTAG GAGCTATAGC
4500

CCATGAAAAA AATGGAAAAC AGATTAATAC GGGTATCGTT GACGATGATG GTATGCTTTA
4560

TATGTCTGGA TTATCAGGGA CAGGGATTAT TAATGTAACA TGAATGGAA AAGTCTGTTC
4620

ATTCCTTTT TCAGAAAAAG ATATATCTAG CAAACAATTA TCTGTTGTAA ATAAACAATG
4680

TTAGGTAGTG CATCCAATTA GTAGAACATG TGTTTTTCGA TAAACGCTCC GATCTCTTTT
4740

TCGTGGATCT CAACTGAGCG TGAGAAGCAG ATTGTTTTAC GAGCCAACCG CTTAATGCGG
4800

GTGCGTAGCG TCAGATTATT ACGCTCAATG CGTTGGGTGA ATATTTTGCC GGTCAGATGC
4860

TTATTCTTCG GTACCC
4876

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: E. coli
- (B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGCCGCATTA TTGCGGCC
18

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: E. coli
 - (B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGCCGCATTA TTGATTGCGG CC
22

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 527 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: E. coli
 - (B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTGACACATT ACGAATGTTA TGTATACAAT AAAAATGATT ATAGCAATAT TAATGGTGTT
60

ATATGAAGAA AACAAATTGGT TTAATTCTAA TTCTTGCTTC ATTCGGCAGC CATGCCAGAA
120

CAGAAATAGC GACTAAAAAC TTCCCAGTAT CAACGACTAT TTCAAAAAGT TTTTTTGCAC
180

CTGAACCACG AATACAGCCT TCTTTTGGTG AAAATGTTGG AAAGGAAGGA GCTTTATTAT
240

TTAGTGTGAA CTTAACTGTT CCTGAAAATG TATCCCAGGT AACGGTCTAC CCTGTTTATG
300

ATGAAGATTA TGGGTTAGGA CGACTAGTAA ATACCGCTGA TGCTTCCCAA TCAATAATCT
360

ACCAGATTGT TGATGAGAAA GGGAAAAAAA TGTTAAAAGA TCATGGTGCA GAGGTTACAC
420

CTAATCAACA AATAACTTTT AAAGCGCTGA ATTATACTAG CGGGGAAAAA AAAATATCTC
480

CTGGAATATA TAACGATCAG GTTATGGTTG GTTACTATGT AACTAA
527

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: E. coli
- (B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

	Met	Lys	Lys	Thr	Ile	Gly	Leu	Ile	Leu	Ile	Leu	Ala	Ser	Phe	Gly
Ser															
	1			5				10						15	

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: E. coli

(B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ser	Met	Leu	Lys	Lys	Ile	Ile	Ser	Ala	Ile	Ala	Leu	Ile	Ala	Gly	Thr
	1				5					10					15
Asn	Gly	Val	Val	Asn	Ala	Gly	Asn	Trp	Gln	Tyr	Lys	Ser	Leu	Asp	Val
				20				25					30		
Arg	Val	Asn	Ile	Glu	Gln	Asn	Phe	Ile	Pro	Asp	Ile	Asp	Ser	Ala	Val
				35				40					45		
Leu	Ile	Ile	Pro	Val	Asn	Tyr	Asp	Ser	Asp	Pro	Lys	Leu	Asp	Ser	Gln
				50				55				60			
Ile	Tyr	Thr	Val	Glu	Met	Thr	Ile	Pro	Ala	Gly	Val	Ser	Ala	Val	Lys
	65						70				75				80
Leu	Ala	Pro	Thr	Asp	Ser	Leu	Thr	Ser	Ser	Gly	Gln	Gln	Ile	Gly	Lys
						85				90				95	
Lys	Val	Asn	Val	Asn	Asn	Pro	Asp	Gln	Asn	Met	Asn	Tyr	Tyr	Ile	Arg
						100				105				110	
Pro	Asp	Ser	Gly	Ala	Gly	Asn	Phe	Met	Ala	Gly	Gln	Lys	Gly	Ser	Phe
						115				120				125	

Arg	Ile Ile Lys Thr Thr Ile Asn Leu Pro Asp Ser Gln Glu Ser Met	65	70	75	80
Ile	Trp Leu Cys Ile Glu Ser Met Pro Pro Ile Glu Lys Ser Thr Lys	85	90	95	
Gly	Asn Arg Lys Glu Gly Arg Thr Asp Ser Ile Asn Ile Ser Ile Arg	100	105	110	
Phe	Cys Ile Lys Leu Ile Tyr Arg Pro Ala Ser Val Pro Ser Pro Val	115	120	125	
Leu	Asn Asn Ile Val Glu Lys Leu Lys Trp His Lys Asn Gly Lys Tyr	130	135	140	
Phe	Val Leu Lys Asn Asn Thr Pro Tyr Tyr Ile Ser Phe Ser Glu Val	145	150	155	
Lys	Phe Asp Ser Asp Lys Val Asn Asn Ala Lys Asp Ile Leu Tyr Val	165	170	175	
Lys	Pro Tyr Ser Glu Lys Lys Ile Asp Ile Ser Asn Arg Ile Ile Lys	180	185	190	
Tyr	Ile Lys Trp Ala Met Ile Asp Asp Ala Gly Ala Lys Thr Lys Leu	195	200	205	
	Glu Ser Ile Leu	210			

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 802 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: E. coli

(B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gln	Met	Asn	Gln	Phe	Tyr	Lys	Lys	Ser	His	Tyr	Ser	Ile	Gln	Lys	His
	1			5				10					15		
Ser	Ile	Thr	Gly	Leu	Leu	Phe	Leu	Leu	Phe	Ile	Tyr	Pro	Phe	Ser	Thr
			20					25					30		
Gly	Tyr	Gly	Asn	Glu	Gln	Phe	Ser	Phe	Asp	Ser	Arg	Phe	Leu	Pro	Ser
			35					40					45		
Leu	Tyr	Asn	Tyr	Ser	Leu	Asn	Ser	Asn	Leu	Pro	Pro	Glu	Gly	Glu	Tyr
		50					55					60			
Pro	Val	Asp	Ile	Tyr	Ile	Asn	Lys	Ile	Lys	Lys	Glu	Ser	Ala	Ile	Ile
	65					70				75				80	
Lys	Phe	Tyr	Ile	Lys	Gly	Asn	Lys	Leu	Val	Pro	Cys	Leu	Ser	Lys	Glu
				85				90						95	
Cys	Ile	Ser	Ser	Leu	Gly	Ile	Asn	Ile	Asn	Asn	Asn	Asp	Asn	Thr	Glu
				100				105						110	

Ser	Val Glu Thr Ser Lys Ala Gly Ile Ser Asn Ile Ser Phe Glu Phe		
	115	120	125
Ile	Ser Leu Arg Leu Phe Ile Ala Val Pro Lys Asn Leu Leu Ser Glu		
	130	135	140
Phe	Asp Lys Ile Ser Ser Lys Asp Ile Asp Asn Gly Ile His Ala Leu		
160	145	150	155
Tyr	Phe Asn Tyr Gln Val Asn Thr Arg Leu Ala Asn Asn Lys Asn Arg		
	165	170	175
Leu	Asp Tyr Ile Ser Val Ser Pro Asn Ile Asn Tyr Phe Ser Trp Arg		
	180	185	190
Glu	Arg Asn Leu Phe Glu Phe Asn Gln Asn Asn Asp Glu Lys Thr Trp		
	195	200	205
Asn	Arg Asn Tyr Thr Tyr Leu Glu Lys Ser Phe Tyr Asp Lys Lys Leu		
	210	215	220
Tyr	Leu Val Val Gly Glu Ser Tyr Thr Asn Ser Asn Val Tyr Asn Asn		
240	225	230	235
Pro	Ser Phe Thr Gly Ile Ser Val Ser Thr Asp Thr Asp Met Tyr Thr		
	245	250	255
Asp	Ser Glu Ile Asp Tyr Thr Pro Glu Ile His Gly Val Ala Asp Ser		
	260	265	270
Ser	Ser Gln Ile Ile Val Arg Gln Gly Asn Thr Ile Ile Ile Asn Glu		

	275	280	285
Thr	Val Pro Ala Gly Pro Phe Ser Phe Pro Ile Thr Asn Leu Met Tyr		
	290	295	300
Lys	Gly Gly Gln Leu Asn Val Glu Ile Thr Asp Ile Tyr Gly Asn Lys		
320	305	310	315
Gly	Gln Tyr Thr Val Asn Asn Ser Ser Leu Pro Val Met Arg Lys Ala		
		325	330 335
Ser	Leu Met Val Tyr Asn Phe Ile Ser Gly Lys Leu Thr Lys Lys Asn		
		340	345 350
His	Glu Asp Gly Asp Phe Phe Thr Gln Gly Asp Ile Asn Tyr Gly Thr		
		355	360 365
Phe	Tyr Asn Ser Thr Leu Phe Gly Gly Tyr Gln Phe Ser Lys Asn Tyr		
		370	375 380
Trp	Asn Leu Ser Thr Gly Ile Gly Thr Asp Leu Gly Phe Ser Gly Ala		
400	385	390	395
Asn	Leu Leu His Val Ser Arg Ser Asn Phe Lys Asn Lys Asn Gly Tyr		
		405	410 415
Val	Ile Asn Leu Gln Gln Asn Thr Gln Leu Arg Pro Phe Asn Ala Gly		
		420	425 430
Asp	Asn Phe Asp Tyr Ala Tyr Arg Lys Lys Arg Tyr Val Glu Leu Ser		
		435	440 445
	Ile Gly Trp His Gly Asn Leu Tyr Asn Gln Leu Lys Asn Ser Phe		

Ser	450	455	460
Asp	Leu Ser Leu Ser Lys Ser Leu Asn Lys Tyr Gly Asn Phe Ser Leu		
480	465	470	475
Met	Tyr Asn Lys Met Lys Tyr Trp Asp Asn Ala Tyr Asp Ser Asn Ser		
	485	490	495
Asn	Ser Ile Arg Tyr Phe Phe Lys Phe Met Arg Ala Met Ile Thr Thr		
	500	505	510
Phe	Cys Ser Leu Asn Lys Tyr Gln Ser Tyr Glu Lys Lys Asp Lys Arg		
	515	520	525
Ser	Ser Ile Asn Ile Ser Leu Pro Leu Thr Lys Asp Tyr Gly His Ile		
	530	535	540
Ser	Ser Asn Tyr Ser Phe Ser Asn Ala Asn Thr Gly Thr Ala Thr Ser		
560	545	550	555
Asn	Val Gly Leu Asn Gly Ser Phe Phe Asn Asp Ala Arg Leu Asn Trp		
	565	570	575
Thr	Ile Gln Gln Asn Arg Thr Thr Arg Asn Asn Gly Tyr Thr Asp Asn		
	580	585	590
Gly	Ser Tyr Ile Ala Thr Ser Tyr Ala Ser Pro Tyr Gly Val Phe Thr		
	595	600	605
Ser	Ser Tyr Ser Gly Ser Asn Lys Tyr Ser Ser Gln Phe Tyr Ser Ala		
	610	615	620

Ala	Gly Gly Ile Val Leu His Ser Asp Gly Val Ala Phe Thr Gln Lys			
640	625	630	635	
Lys	Gly Asp Thr Ser Ala Leu Val Arg Ile Asp Asn Ile Ser Asp Ile			
	645	650	655	
Leu	Ile Gly Asn Thr Pro Gly Val Tyr Thr Gly Tyr Asn Gly Phe Ala			
	660	665	670	
Asp	Ile Pro His Leu Gln Pro Phe Lys Lys Asn Thr Ile Leu Ile Asn			
	675	680	685	
Val	Lys Gly Ile Pro Asp Gly Ile Thr Leu Ala Asn Ile Lys Lys Gln			
	690	695	700	
Lys	Ile Pro Ser Arg Gly Ala Ile Val Lys Val Lys Phe Asp Ala Lys			
720	705	710	715	
Pro	Gly Asn Asp Ile Leu Phe Lys Leu Thr Thr Lys Asp Gly Lys Thr			
	725	730	735	
Thr	Pro Leu Gly Ala Ile Ala His Glu Lys Asn Gly Lys Gln Ile Asn			
	740	745	750	
Gly	Gly Ile Val Asp Asp Asp Gly Met Leu Tyr Met Ser Gly Leu Ser			
	755	760	765	
Pro	Thr Gly Ile Ile Asn Val Thr Trp Asn Gly Lys Val Cys Ser Phe			
	770	775	780	
Lys	Phe Ser Glu Lys Asp Ile Ser Ser Lys Gln Leu Ser Val Val Asn			
800	785	790	795	